

# UCSF

## UC San Francisco Previously Published Works

### Title

Radical Prostatectomy Versus Radiation and Androgen Deprivation Therapy for Clinically Localized Prostate Cancer: How Good Is the Evidence?

### Permalink

<https://escholarship.org/uc/item/89z8z1tc>

### Journal

International Journal of Radiation Oncology • Biology • Physics, 93(5)

### ISSN

0360-3016

### Authors

Roach, Mack  
Lizarraga, Tania L Ceron  
Lazar, Ann A

### Publication Date

2015-12-01

### DOI

10.1016/j.ijrobp.2015.08.005

Peer reviewed

Critical Review

# Radical Prostatectomy Versus Radiation and Androgen Deprivation Therapy for Clinically Localized Prostate Cancer: How Good Is the Evidence?



Mack Roach, III, MD,\* Tania L. Ceron Lizarraga, MD,\*  
and Ann A. Lazar, PhD\*<sup>†</sup>

*\*Department of Radiation Oncology and <sup>†</sup>Division of Oral Epidemiology and Dental Public Health, Division of Biostatistics, University of California San Francisco, San Francisco, California*

Received Jul 7, 2015, and in revised form Jul 29, 2015. Accepted for publication Aug 3, 2015.

## Purpose

The optimal treatment of clinically localized prostate cancer is controversial. Most studies focus on biochemical (PSA) failure when comparing radical prostatectomy (RP) with radiation therapy (RT), but this endpoint has not been validated as predictive of overall survival (OS) or cause-specific survival (CSS). We analyzed the available literature to determine whether reliable conclusions could be made concerning the effectiveness of RP compared with RT with or without androgen deprivation therapy (ADT), assuming current treatment standards.

## Methods

Articles published between February 29, 2004, and March 1, 2015, that compared OS and CSS after RP or RT with or without ADT were included. Because the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system emphasis is on randomized controlled clinical trials, a reliability score (RS) was explored to further understand the issues associated with the study quality of observational studies, including appropriateness of treatment, source of data, clinical characteristics, and comorbidity. Lower RS values indicated lower reliability.

## Results

Fourteen studies were identified, and 13 were completely evaluable. Thirteen of the 14 studies (93%) were observational studies with low-quality evidence. The median RS was 12 (range, 5-18); the median difference in 10-year OS and CSS favored RP over RT: 10% and 4%, respectively. In studies with a  $RS \leq 12$  (average RS 9) the 10-year OS and CSS median differences were 17% and 6%, respectively. For studies with a  $RS > 12$  (average RS 15.5), the 10-year OS and CSS median differences were 5.5% and 1%, respectively. Thus, we observed an association between low RS and a higher percentage difference in OS and CSS.

## Conclusions

Reliable evidence that RP provides a superior CSS to RT with ADT is lacking. The most reliable studies suggest that the differences in 10-year CSS between RP and RT are small, possibly  $< 1\%$ . © 2015 Elsevier Inc. All rights reserved.

Reprint requests to: Mack Roach III, MD, Department of Radiation Oncology, University of California San Francisco, Helen Diller Family Comprehensive Cancer Center, 1600 Divisadero St, Suite H 1031, San

Francisco, CA 94143-1708. Tel: (415) 353-7181; E-mail: [mroach@radonc.ucsf.edu](mailto:mroach@radonc.ucsf.edu)

Conflict of interest: none.

## Introduction

The management of clinically localized prostate cancer is controversial. Recent evidence suggests that men with low-risk disease may be optimally managed with active surveillance, whereas those with intermediate-risk and high-risk disease appear to require definitive treatment (1). Some studies focusing on biochemical (prostate-specific antigen [PSA]) control suggest that treatments based on radiation therapy (RT) might be as good as, or better than, radical prostatectomy (RP) (2). However, different PSA endpoints are used for these therapies, and PSA failure is not accepted as a surrogate endpoint for overall survival (OS) or cause-specific survival (CSS) (3). The preferred treatment would best be identified by large phase 3 randomized trials. Unfortunately, limited data from small randomized trials of dubious relevance have failed to resolve this question (1). To address this issue, several studies led by investigators from the urologic community have been reported, with some concluding that RP renders a better chance of OS and CSS than RT (4). However, most of these studies did not compare RP with treatment that was consistent with standard of care (eg, in accordance with National Comprehensive Cancer Center [NCCN] guidelines) (5, 6). There is clear level I evidence derived from phase 3 trials demonstrating that the inclusion of short-term androgen deprivation therapy (ADT) substantially improves OS in patients with intermediate-risk disease (1). There is also clear level I evidence that long-term ADT combined with RT substantially improves OS compared with RT alone or RT with short-term ADT in men with high-risk prostate cancer (1). Furthermore, most retrospective studies comparing RP with RT show considerable bias in patient selection, with major differences in the baseline characteristics between cohorts (4). In addition, some studies included several thousand patients and demonstrated very small but statistically significant differences in OS, which may not be clinically significant. Our semi-quantitative analysis was undertaken to specifically apply metrics to assess the reliability of the conclusions in these studies, when taking into account the contemporary treatment standards for RT.

## Methods and Materials

### Study selection criteria

The studies included in the analysis met the following requirements: (1) published in the past 10 years (February 29, 2004, through March 1, 2015); (2) compared 10-year OS and CSS in patients treated with RP or RT with or without ADT; (3) provided sufficient data to allow an analysis of pretreatment prognostic factors in cohorts of patients without metastatic disease; (4) were identified by performing a literature search in PubMed using “radical prostatectomy versus radiation,” “randomized trials,” and/

or was included in a recently published meta-analysis (which identified only 7 articles) and/or was cited in other recent studies comparing RP with RT (5). Studies reporting only CSS after RP or RT were excluded from this analysis because they precluded an assessment of inherent differences in the patient selection as might be reflected in OS in the cohorts being compared (7, 8).

### GRADE approach

An attempt was made to evaluate the quality of the evidence provided in these studies using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system (9, 10). The Cochrane Collaboration adopted the principles of the GRADE system for evaluating the quality of evidence described in reviews. The GRADE system specifies 4 levels (high, moderate, low, or very low) of quality of evidence; the highest quality rating is designated for randomized clinical trial (RCT) evidence, and the lowest quality rating is designated for observational studies. Inasmuch as the majority of the eligible studies were observational in this review, rather than RCTs, these observational studies were considered to be of low quality (11). Because the GRADE system emphasis is based on RCTs, a reliability scale was explored to assess the issues associated with the study quality of these observational studies.

### Reliability score and its rationale

We constructed a reliability scale in a fashion similar to that in previously reported studies (12-14). The basic steps involved assigning reliability points from 0 to 5, based on our perceptions about the relative importance of each factor, inasmuch as they might be expected to have an impact on OS and CSS. Each study was independently reviewed by 2 authors, and a composite reliability score (RS) was assigned based on a sum of the factors, after a consensus was reached among all the reviewers. The RSs assigned are summarized in Table 1, where higher scores indicate better reliability. For reflecting the quality of the source of data, 5 points were assigned for phase 3 randomized trials, and 4 points were assigned for prospective 1-institution or 2-institution studies because these types of studies insure the greatest uniformity of quality of care. By contrast, only 1 point was assigned to Surveillance Epidemiology and End Results (SEER) data because before 2010, biopsy Gleason score was available only for RT patients, whereas pathologic GS was available for RP patients. Early studies also did not use the TNM staging, and recent studies have demonstrated reasons for the cautious use of these data (15).

For assessing treatment, 5 and 4 points, respectively, were assigned for studies that included data based on patients treated with RT in accordance with NCCN guidelines and level I evidence, such that high-risk patients received

**Table 1** Reliability score: endpoint overall survival

1. Source of data	
a. Randomized trial yes = 5	
b. Matched single-institution or 2-institution data = 3	
c. Nonrandomized observational databases = 2	
d. Retrospective based such as SEER = 1	
e. Other	
2. Appropriate use ADT (none for low-risk, short-term ADT + RT for intermediate-risk, long-term ADT for high-risk prostate cancer)*	
a. RT + long-term ADT <sup>†</sup> all high-risk = 5	
b. RT + short-term ADT <sup>‡</sup> all intermediate-risk = 4	
c. Mixed seeming appropriate but not explicit = 3	
d. ADT in minority perhaps appropriately = 2	
e. ADT omitted in many for whom it was indicated = 1	
f. RT alone = 0	
3. Charlson comorbidity index	
a. Yes = 2	
b. No = 0	
4. Additional comorbidity adjustments (non—prostate cancer—related causes)	
a. Detailed = 2	
b. Relevant data (less detailed) = 1	
c. Minor information or none = 0	
5. Sample size (not too big or small)	
a. Clinically relevant = 2	
b. Very large = 1 (n>12,000)	
c. Too small = 0 (n<400)	
6. Survival curves adjusted for prognostic factors	
a. Yes = 2	
b. Partial = 1	
c. No = 0	
7. Provided information to assess cancer-related prognostic factors including:	
a. Gleason score all patients (biopsy only) = 2	
b. Gleason score vast majority of patients = 1	
c. Gleason score less than majority = 0	
d. T stage	
I. All = 2	
II. Vast majority or unclear how used = 1	
III. No = 0	
e. PSA	
I. All = 2	
II. Vast majority = 1	
III. None = 0	

*Abbreviations:* ADT = androgen deprivation therapy; PSA = prostate-specific antigen; RT = external beam radiation therapy; SEER = Surveillance Epidemiology and End Results.

\* Risk determined according to National Comprehensive Cancer Center guidelines (6).

<sup>†</sup> Long-term ADT = 2-3 years or more.

<sup>‡</sup> Short-term ADT = 4-6 months.

long-term ADT with RT and intermediate-risk patients received short-term ADT. Studies including high-risk and intermediate-risk patients who were treated exclusively with RT alone were assigned a score of 0 because conclusions about such patients would not be relevant to our current standard of care. Studies including Charlson or similar comorbidity data, or making other significant evidence-based adjustments for comorbidity, were credited with up to 2 points. Although the former are fraught with limitations, inclusion of these data does allow a reduction in biases compared with no such adjustment (15).

Inasmuch as most phase 3 randomized trials with an OS endpoint included between 450 and 2000 patients, we considered the optimal study size to be between 400 and 12,000 patients (1). For example, in the landmark trial comparing RP with “watchful waiting,” nearly 700 patients were required to show a small OS advantage, when a no-treatment control arm was used (16). Therefore, it is likely that when comparing 2 effective treatments, a substantially larger study might be required. Thus, studies that included fewer than 400 patients were considered to be less reliable and were awarded 1 point. By contrast, studies

including more than 12,000 patients were also considered less reliable because even small degrees of selection or attribution bias could translate into statistically significant differences in survival that might not be clinically relevant, potentially rendering false positive results. Some studies provided unadjusted or partially adjusted OS and CSS survival curves that tended to distort the apparent differences between RP and RT; accordingly, these were marked down by 1 or 2 points.

Finally, points were also assigned for the degree to which GS, T stage, and PSA were accounted for (Table 1). We based our analysis on studies reporting 10-year results because very few events occur before 10 years. In addition, studies published before 10 years ago, with such follow-up, would have been based on patients whose treatment occurred up to 20 years ago, when the current standards of treatment were not yet established. To assess the overall treatment differences in 10-year OS and 10-year CSS, linear mixed regression models of 10-year OS and 10-year CSS (outcomes) were generated separately with a fixed treatment (RP vs RT) covariate and random intercepts to account for variability across studies (17). An interaction term between treatment and RS was used to explore potential heterogeneity of the treatment effects according to RS. The RS was dichotomized, with low RS corresponding to studies to  $\leq 12$ , and high RS corresponding to studies with  $RS > 12$ . This cutoff is the median of the distribution of the reliability score among all eligible studies. All evaluable studies were included in the statistical analysis. All statistical tests provided 2-sided *P* values, and *P* values  $< .05$  were considered statistically significant. Statistical analyses were performed with SAS version 9.4.

## Results

Using the selection criteria just described, a total of 14 studies were identified for inclusion in this analysis (4, 18-29). One of them was subsequently noted to be only partially analyzable because we were unable to determine the CSS (29). Thus, 13 studies were completely evaluable for OS and CSS. One was a randomized controlled clinical trial, and the remaining were observational studies. Based on the GRADE system, 93% (13 observational of 14) of the studies were considered to have a low level of confidence in quality (4, 19-23, 24-30). Table 2 lists the 14 studies in chronologic order, the RS assigned to each, and the scoring of each individual component. The median and average RS for all studies was 12 (range, 5-18), and the overall median difference in OS and CSS favoring RP over RT was 10% (average difference 11%; 95% confidence interval [CI]: 7%-16%) and 4% (average difference 5%; 95% CI: 2%-8%) at 10 years, respectively. For most studies (80%), OS was statistically significant, but CSS was not (46%). To evaluate OS, all the studies reported using Kaplan-Meier analysis and Cox proportional hazards models. For CSS, 5 (38%) of the studies also reported using Kaplan-Meier analysis and Cox models,

and the remaining reported using cumulative incidence measures, including the Fine-Gray subdistribution hazard model to account for competing risks, such as deaths of other causes. Nearly all (93%) the studies reported older men treated with RT rather than with RP (average age RT: 69 vs RP: 64), and they were more likely (80%) to have statistically significantly higher baseline tumor grades and more comorbidities than were those treated with RP. Half (7/14) of the studies did not report the total number of deaths or deaths of prostate cancer by treatment status.

The lowest RS of 5 was given to study 4, which received 1 point for SEER data, 0 points for RT without ADT for intermediate-risk or high-risk disease, no comorbidity data available, very small sample size  $\leq 400$  patients, no survival curve adjustments, 1 point for GS in 93% of patients, 2 points for having T stage in all patients, and 1 point for having PSA in 85% of patients. By contrast, study 8 received 3 points for being based on 2 institutions (supporting a uniformity of care) and 3 points for treatment because 34% of patients treated with external beam RT or brachytherapy received neoadjuvant, concurrent, or adjuvant ADT, including 45% with intermediate-risk and 82% with high-risk disease with a median of 6 months. This study also received 2 points for Charlson comorbidity index; 2 points for size (including 5811 patients); 2 points for survival curve adjustments; and 2 points for GS, T stage, and PSA data available respectively, adding to a total of 18 points.

Figure 1 provides graphic summaries of the relationship between RS and the median differences between OS and CSS in patients treated with RP compared with RT. Above 0 indicates benefit to RP, and below 0 indicates benefit to RT. The lower the RS, the larger the OS and CSS benefit to RP appeared to be. Those studies with a  $RS \leq 12$  (average RS 9) suggested a 10-year OS median difference of 17% (average difference of 17%, 95% CI: 13%-20%;  $P < .0001$ ) and CSS difference of 6% (average difference of 7%, 95% CI: 4%-11%;  $P = .0009$ ). By contrast, for those studies with a  $RS > 12$  (average RS 15.5), the median difference in 10-year OS was 5.5% (average difference of 5%, 95% CI: 1%-9%;  $P = .01$ ) and the median difference in CSS was 1% (average difference of 2%, 95% CI: -2% to 7%;  $P = .24$ ). Thus, the 10-year difference in OS between low and high reliability studies was 12% (95% CI: 7%-16%;  $P = .0002$ ) and for CSS was 5% (-1% to 10%;  $P = .10$ ).

## Discussion

In this analysis we applied a set of assumptions to resolve divergent conclusions concerning the effectiveness of RP compared with RT for localized prostate cancer. We focused on OS and CSS because recent studies have advocated RP as the treatment of choice but were compromised by biases (31). Previous studies have addressed PSA control rates, but PSA failure is not an established surrogate for OS or CSS.

**Table 2** Studies included in the analysis

No.	Study	Source of data*	Use of ADT†	CCI	Other adjustments‡	Sample size	Survival curves§	Prognostic factors	Total score
1	Akakura et al (18)	5	5	0	0	0	0	6	16
2	Tward et al (19)	1	3	2	2	1	0	4	13
3	Tewari et al (20)	3	0	2	1	0	0	3	9
4	Merglen et al (21)	1	0	0	0	0	0	4	5
5	Liu et al (22)	1	0	2	0	2	0	3	8
6	Boorjian et al (23)	3	3	2	0	2	2	6	18
7	Abdollah et al (30)	1	0	2	0	1	1	4	9
8	Kibel (24)	3	3	2	0	2	2	6	18
9	Sooriakumaran (4)	1	0	2	0	1	0	6	10
10	Hoffman et al (25)	1	1	2	0	2	0	6	12*
11	Nepple et al (26)	3	0	2	1	2	0	6	14
12	Yamamoto et al (27)	3	1	2	0	0	0	6	12
13	Sun et al (28)	1	0	2	0	1	0	4	8
14	Tsao et al (29)	3	3	0	0	2	0	6	14

Abbreviations: ADT = androgen deprivation therapy; CCI = Charlson comorbidity index; PSA = prostate-specific antigen; RT = radiation therapy; SEER = Surveillance Epidemiology and End Results.

\* Randomized versus single institution versus SEER.

† Appropriate if long-term ADT with RT for high risk and short-term ADT for intermediate risk according to National Comprehensive Cancer Center guidelines (6).

‡ Additional comorbidity adjustments not included in Charlson index not related to prostate cancer.

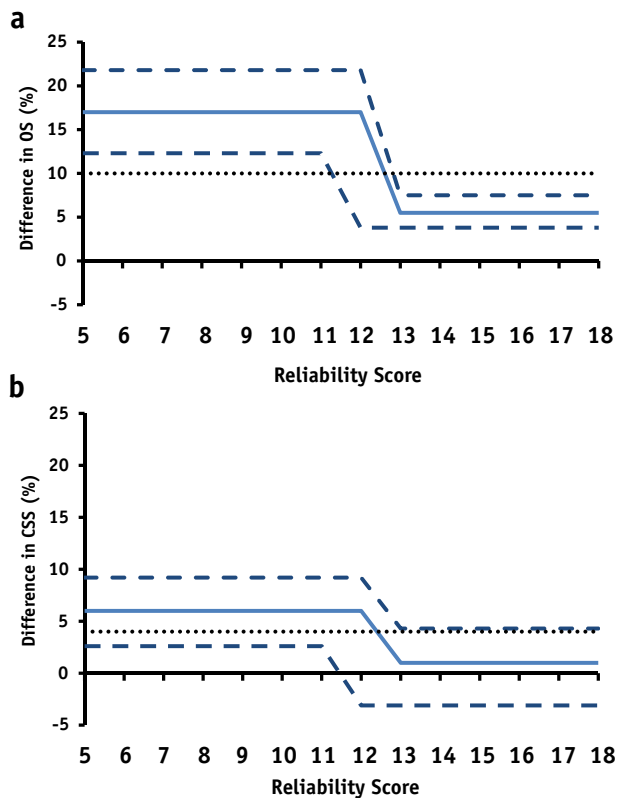
§ Adjusted survival curves versus partially versus not at all.

|| Gleason score, T stage, PSA.

We observed that based on the most reliable studies, patients who undergo RP appear to have a 10-year OS 5% higher than those who have been given appropriately administered RT and a CSS advantage of approximately 1%. We conclude that the residual OS differences most likely reflect differences resulting from the fact that many studies provide unadjusted survival curves, and there remain other unaccounted-for differences between patients who undergo RP and those who receive RT. For example, Giordano et al (15) showed that using SEER data, major residual biases persisted, creating “improbable results.” These investigators concluded that “Controlling for comorbidity, extent of disease, and other characteristics by multivariate analyses or by propensity analyses had remarkably small impact on these improbable results” and that “the results from observational studies of treatment outcomes should be viewed with caution.” It was in part based on these data that we decided to give limited value to the Charlson comorbidity index. This study and several studies suggest that adjustments based on the Charlson comorbidity index failed to insure adequate adjustments and were best applied when populations were more homogeneous, such as those limited to a Charlson comorbidity score of 0 (26). The inherent OS biases favoring patients treated by RP were highlighted in the report by Eifer et al (32), who showed that men undergoing RP had a substantially lower risk of all-cause mortality than expected. They noted that “overall death rate was lower in men treated with radical prostatectomy than in the general American population (standardized mortality ratio 0.47, 95% CI 0.44-0.49).”

The residual 1% CSS differences favoring RP over RT might also well be explained by the fact the included studies did not adjust for other factors associated with a worse outcome (8). In addition, several classified a death as “due to prostate cancer” if this diagnosis was listed on any of the top 3 lines of the death certificate (20). Because patients treated with RT were more likely to die of other causes, they would be at greater risk for attribution bias, which could certainly raise their apparent risk of CSS by 1% or more. It would also be important for patients to balance an excess 1% difference in CSS at 10 years against the higher rate of urinary incontinence, erectile dysfunction, and the higher 30-day, 60-day, and 90-day mortality rate associated with RP (33).

A major criticism likely to be rendered by the urologic community is the reliability scale. Given that nearly all the eligible studies were observational, the GRADE approach would initially consider nearly all these studies as providing a low quality of evidence. We would argue that our analysis systematically takes on a whole host of biases not accounted for by any of the studies cited. We readily admit that our assessment of reliability approach is far from perfect, as further detailed below, but none of the current validated instruments, such as GRADE, could be readily applied to adequately address these issues either, and none of these approaches would help with estimating the differences in OS at 10 years (9, 10). Although these approaches, such as GRADE, may not recommend estimating the differences in OS at 10 years because the biases of such observational studies could not be addressed, we believe that our exploratory analyses provide clinically relevant



**Fig. 1.** Graphical summary of the median difference in the estimated 10-year overall survival (OS) (a) and cause-specific survival (CSS) (b) by the dichotomized reliability score ( $\leq 12$  vs  $> 12$ ). Above 0 indicates benefit for radical prostatectomy, 0 no differences between treatments, and below 0 a benefit to radiation therapy with or without androgen deprivation therapy. Broken lines (—) indicate the pointwise 95% confidence interval based on bootstrap estimation of 100,000 replicates. Black dotted line indicates the overall differences.

information. These estimates are based on peer-reviewed published data and provide information on the order of the magnitude of the reported differences in survival. We believe they are more helpful than no estimate at all. We also present the individual components of the scale and the raw outcome data so readers can draw their own conclusions.

Several weaknesses of our study design are worthy of comment. First, we had to decide how much to reduce factors by, to account for shortcomings. For example, we tended to place a relatively low value on studies based on SEER data because before 2010, tumor grade was assigned based on the pathologic (postoperative) grade for RP patients but the biopsy grade for RT patients, thus creating a bias in the most important predictor of death in men with clinically localized disease, as described above. Second, although several of these studies reported using propensity adjustments to account for the differences between patients treated with RP and RT, we did not weight this heavily. There were 2 major reasons for doing so: (1) they

did not provide enough information for us to assess the magnitude of the reliability credit (if any) that should be assigned to them; and (2) because they did not provide adjusted OS curves, we used their unadjusted curves. If we had also given them credit for the propensity adjustments and used these unadjusted estimates, we would have artificially inflated the OS difference estimate and possibly their RS. Regarding the first point, there is a major lack of clarity as to how the authors accounted for large differences in the study populations. For example, in 1 study, 75% of patients undergoing RP were  $< 67$  years of age, and 75% of men treated with RT were  $> 67$  years of age (8). Because the benefits of RP compared with no initial treatment were largely limited to men  $< 65$  years of age (in a landmark phase 3 trial), it is unclear how one would adjust with confidence for the expected outcomes when  $\leq 25\%$  of patients treated with RT might be expected to have a favorable outcome if they had been treated with RP (8). Finally, the studies claiming to be propensity adjusted generally had low reliability scores, so that adding a couple of points would not change our conclusions (data not shown).

Another limitation of our study was that we based our reliability score primarily on how RT was delivered. Because there is no level I evidence (that we are aware of) suggesting that changes in RP technique have improved OS or CSS, we focused on crediting studies based on the degree to which ADT was appropriately added to RT (1). Inasmuch as neither dose-escalated RT, whole pelvic irradiation, 3-dimensional conformal RT, intensity modulated RT, or brachytherapy have been shown to prolong OS or CSS, we declined to include these factors in our RS, although some of these advances might improve the results seen with RT with or without ADT (1, 34, 35).

Another limitation is that we compared the degree to which RP was superior to RT. It is conceivable that the results with RT are superior to those with RP for men with high-risk locally advanced prostate cancer (36). To date the level I evidence for the benefits of RP compared with no initial treatment (observation or watchful waiting) seemed to be limited to men  $< 65$  years of age. By contrast, there is a larger body of evidence that RT combined with ADT improves OS and CSS compared with RT alone or ADT alone, even in men with a median age of 70 years (1). Thus, one could argue that the burden of proof should lie with those seeking to argue for RP in men with high-risk disease, especially if they are over 65.

The strength of this study is that it is based on all the articles we could identify comparing RP with RT reporting OS and CSS for men with clinically localized prostate cancer. We compared RT-based treatments according to level I evidence-based treatment to RP and focused on OS and CSS. To our knowledge, this is the most detailed assessment of this topic, and although it does not resolve the controversy, it highlights why to date there remains no consensus about the best treatment for these patients. We predict that a fairly large phase 3 trial would be required to

resolve this question, and it is doubtful that such a trial will occur in the immediate future.

## References

- Roach M 3rd, Thomas K. Overview of randomized controlled treatment trials for clinically localized prostate cancer: Implications for active surveillance and the United States preventative task force report on screening? *J Natl Cancer Inst Monogr* 2012;2012:221-229.
- Grimm P, Billiet I, Bostwick D, et al. Comparative analysis of prostate-specific antigen free survival outcomes for patients with low, intermediate and high risk prostate cancer treatment by radical therapy. Results from the Prostate Cancer Results Study Group. *BJU Int* 2012;109(Suppl 1):22-29.
- Lee BH, Kibel AS, Ciezki JP, et al. Are biochemical recurrence outcomes similar after radical prostatectomy and radiation therapy? Analysis of prostate cancer-specific mortality by nomogram-predicted risks of biochemical recurrence. *Eur Urol* 2015;67:204-209.
- Sooriakumaran P, Nyberg T, Akre O, et al. Comparative effectiveness of radical prostatectomy and radiotherapy in prostate cancer: Observational study of mortality outcomes. *BMJ* 2014;348:g1502.
- Petrelli F, Vavassori I, Coiru A, et al. Radical prostatectomy or radiotherapy in high-risk prostate cancer: A systematic review and metaanalysis. *Clin Genitourin Cancer* 2014;12:215-224.
- Mohler JL, Armstrong AJ, Bahnson RR, et al. Prostate cancer, version 3.2012: Featured updates to the NCCN guidelines. *J Natl Compr Canc Netw* 2012;10:1081-1087.
- Degroot JM, Brundage MD, Lam M, et al. Prostate cancer-specific survival differences in patients treated by radical prostatectomy versus curative radiotherapy. *Can Urol Assoc J* 2013;7:E299-E305.
- Cooperberg MR, Vickers AJ, Broering JM, et al. Comparative risk-adjusted mortality outcomes after primary surgery, radiotherapy, or androgen-deprivation therapy for localized prostate cancer. *Cancer* 2010;116:5226-5234.
- Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches. The GRADE Working Group. *BMC Health Serv Res* 2004;4:38.
- Phi L, Ajaj R, Ramchandani MH, et al. Expanding the grading of recommendations assessment, development, and evaluation (Ex-GRADE) for evidence-based clinical recommendations: Validation study. *Open Dent J* 2012;6:31-40.
- Guyatt GH, Oxman AD, Schunemann HJ. GRADE guidelines: An introduction to the 10th-13th articles in the series. *J Clin Epidemiol* 2013;66:121-123.
- Roach M 3rd, Alexander M. The prognostic significance of race and survival from breast cancer: A model for assessing the reliability of reported survival differences. *J Natl Med Assoc* 1995;87:214-219.
- Bach PB, Schrag D, Brawley OW, et al. Survival of blacks and whites after a cancer diagnosis. *JAMA* 2002;287:2106-2113.
- Rivin del Campo E, Thomas K, Weinberg V, et al. Erectile dysfunction after radiotherapy for prostate cancer: A model assessing the conflicting literature on dose-volume effects. *Int J Impot Res* 2013;25:161-165.
- Giordano SH, Kuo YF, Duan Z, et al. Limits of observational data in determining outcomes from cancer therapy. *Cancer* 2008;112:2456-2466.
- Bill-Axelsson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med* 2014;370:932-942.
- Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;38:963-974.
- Akakura K, Suzuki H, Ichikawa T, et al. A randomized trial comparing radical prostatectomy plus endocrine therapy versus external beam radiotherapy plus endocrine therapy for locally advanced prostate cancer: Results at median follow-up of 102 months. *Jpn J Clin Oncol* 2006;36:789-793.
- Tward JD, Lee CM, Pappas LM, et al. Survival of men with clinically localized prostate cancer treated with prostatectomy, brachytherapy, or no definitive treatment: Impact of age at diagnosis. *Cancer* 2006;107:2392-2400.
- Tewari A, Divine G, Chang P, et al. Long-term survival in men with high grade prostate cancer: A comparison between conservative treatment, radiation therapy and radical prostatectomy: A propensity scoring approach. *J Urol* 2007;177:911-915.
- Merglen A, Schmidlin F, Fioretta G, et al. Short- and long-term mortality with localized prostate cancer. *Arch Intern Med* 2007;167:1944-1950.
- Liu L, Coker AL, Du XL, et al. Long-term survival after radical prostatectomy compared to other treatments in older men with local/regional prostate cancer. *J Surg Oncol* 2008;97:583-591.
- Boorjian SA, Karnes RJ, Viterbo R, et al. Long-term survival after radical prostatectomy versus external-beam radiotherapy for patients with high-risk prostate cancer. *Cancer* 2011;117:2883-2891.
- Kibel AS, Ciezki JP, Klein EA, et al. Survival among men with clinically localized prostate cancer treated with radical prostatectomy or radiation therapy in the prostate specific antigen era. *J Urol* 2012;187:1259-1265.
- Hoffman RM, Koyama T, Fan KH, et al. Mortality after radical prostatectomy or external beam radiotherapy for localized prostate cancer. *J Natl Cancer Inst* 2013;105:711-718.
- Nepple KG, Stephenson AJ, Kallogjeri D, et al. Mortality after prostate cancer treatment with radical prostatectomy, external-beam radiation therapy, or brachytherapy in men without comorbidity. *Eur Urol* 2013;64:372-378.
- Yamamoto S, Kawakami S, Yonese J, et al. Long-term oncological outcome in men with T3 prostate cancer: Radical prostatectomy versus external-beam radiation therapy at a single institution. *Int J Clin Oncol* 2014;19:1085-1091.
- Sun M, Sammon JD, Becker A, et al. Radical prostatectomy vs radiotherapy vs observation among older patients with clinically localized prostate cancer: A comparative effectiveness evaluation. *BJU Int* 2014;113:200-208.
- Tsao CK, Gray KP, Nakabayashi M, et al. Patients with biopsy Gleason 9 and 10 prostate cancer have significantly worse outcomes compared to patients with Gleason 8 disease. *J Urol* 2015;194:91-97.
- Abdollah F, Schmitges J, Sun M, et al. Comparison of mortality outcomes after radical prostatectomy versus radiotherapy in patients with localized prostate cancer: A population-based analysis. *Int J Urol* 2012;19:836-844.
- Roach M 3rd. Radical prostatectomy v radiation: Only a randomised trial can provide the answer. *BMJ* 2014;348:g2266.
- Eifler JB, Humphreys EB, Agro M, et al. Causes of death after radical prostatectomy at a large tertiary center. *J Urol* 2012;188:798-801.
- Hansen J, Gandaglia G, Bianchi M, et al. Re-assessment of 30-, 60- and 90-day mortality rates in non-metastatic prostate cancer patients treated either with radical prostatectomy or radiation therapy. *Can Urol Assoc J* 2014;8:E75-E80.
- Kim MM, Hoffman KE, Levy LB, et al. Prostate cancer-specific mortality after definitive radiation therapy: Who dies of disease? *Eur J Cancer* 2012;48:1664-1671.
- Kalbasi A, Li J, Berman A, et al. Dose-escalated irradiation and overall survival in men with nonmetastatic prostate cancer. *JAMA Oncol* 2015 Jul 16 [Epub ahead of print].
- Arcangeli G, Strigari L, Arcangeli S, et al. Retrospective comparison of external beam radiotherapy and radical prostatectomy in high-risk, clinically localized prostate cancer. *Int J Radiat Oncol Biol Phys* 2009;75:975-982.